

TITLE OF THE INVENTION
COMPUTER PROGRAM PRODUCT, METHOD, AND SYSTEM OF DOCUMENT ANALYSIS

CROSS-REFERENCE TO RELATED APPLICATIONS

5 This application is based upon and claims the benefit of
priority from the prior Japanese Patent Application No.2001-079349,
filed March 19, 2001, the entire contents of which are incorporated
herein by reference.

BACKGROUND OF THE INVENTION

10 1. Field of the Invention

本発明は、文書データの分析作業を支援するコンピュータプログラムプロダクト及び文書分析方法並びに文書分析システムに関する。

2. Description of the Related Art

インターネット、イントラネット、エクストラネットなどの技術向上により、
15 企業内又は企業間で情報の収集、情報の共有化が図られている。
企業は集めた情報に対して各種の分析を行うなどにより情報の有効利用を試みる。

しかし、企業が例えば日報データなどのデータを計算機システムで管理すると、
収集されるデータ数が膨大になり、計算機システムのユーザは、集めた日報データ
20 に含まれている有効な情報を把握することが困難になる場合がある。

また、収集された日報データの記載量が多い場合には、その日報データの中から
有効又は特徴的な記載部分を探すまでのユーザの労力が大きくなる。

したがって、日報データから有効又は特徴的な情報を把握する作業の効率化が
求められる。

25 さらに、収集したデータに含まれている有効又は特徴的な情報をユーザが適切

に把握するためにシステムの操作性の向上が望まれる。

BRIEF SUMMARY OF THE INVENTION

本発明は、大量の文書データを扱う計算機システムにおいて、有効な情報を把握容易とするコンピュータプログラムプロダクト及び文書分析方法並びに文書分析システムを提供することを目的とする。

According to an embodiment of the present invention, there is provided an article of manufacture comprising a computer usable medium having computer readable program code means embodied therein, the computer program code means comprising

10 a computer readable program code that 要約に含めるために抽出する要素として定義された要約要素を含む用語定義辞書データを参照し、分析対象の文書データに含まれている要約要素を抽出する、

a computer readable program code that 抽出された前記要約要素を所定の規則にしたがって組み合わせて前記分析対象の文書データの要約情報を生成する、

15 and

a computer readable program code that 前記分析対象の文書データと前記要約情報をリンクする。

According to a still another embodiment of the present invention, there is provided an article of manufacture comprising a computer usable
20 medium having computer readable program code means embodied therein, the computer program code means comprising

a first computer readable program code that 要約に含めるために抽出する要素として定義された要約要素を含む用語定義辞書データを参照し、分析対象の文書データに含まれている要約要素を抽出する、

25 a second computer readable program code that 抽出された前記要約要素を所定の規則にしたがって組み合わせて前記分析対象の文書データの要約情報を生成する、

a third computer readable program code that 前記分析対象の文書データと前記要約情報をリンクする、 and

a fourth computer readable program code that ユーザから前記要約情報の指定を受け付けた場合に、前記分析対象の文書データと前記要約情報とのリンク結果に基づいて、指定された前記要約情報に対応する前記分析対象の文書データを検索し、指定された前記要約情報と検索された前記分析対象の文書データとを含む画面データを生成する。

According to a still another embodiment of the present invention, there is provided a method of document analysis by a computer, comprising

要約に含めるために抽出する要素として定義された要約要素を含む用語定義辞書データを参照し、

分析対象の文書データに含まれている要約要素を抽出し、

抽出された前記要約要素を所定の規則にしたがって組み合わせて前記分析対象の文書データの要約情報を生成し、

前記分析対象の文書データと前記要約情報をリンクする。

According to a still another embodiment of the present invention, there is provided a method of document analysis by a computer, comprising

要約に含めるために抽出する要素として定義された要約要素を含む用語定義辞書データを参照し、

分析対象の文書データに含まれている要約要素を抽出し、

抽出された前記要約要素を所定の規則にしたがって組み合わせて前記分析対象の文書データの要約情報を生成し、

前記分析対象の文書データと前記要約情報をリンクし、

ユーザから前記要約情報の指定を受け付けた場合に、前記分析対象の文書データと前記要約情報とのリンク結果に基づいて、指定された前記要約情報に対応する前記分析対象の文書データを検索し、

指定された前記要約情報と検索された前記分析対象の文書データとを含む画面データを生成する。

According to a still another embodiment of the present invention,

there is provided a method of document analysis by a computer, comprising

所属するカテゴリを示すインデックス情報を含む分析対象の文書データを受け付け、

- 5 要約に含めるために抽出する要素として定義された要約要素を含む用語定義辞書データを参照し、

前記分析対象の文書データに含まれている要約要素を抽出し、

抽出された前記要約要素を所定の規則にしたがって組み合わせて前記分析対象の文書データの要約情報を生成し、

10 前記分析対象の文書データと前記要約情報をリンクし、

ユーザから前記カテゴリの指定を受け付けた場合に、前記インデックス情報に基づいて、指定された前記カテゴリに所属する前記分析対象の文書データを検索し、

15 前記分析対象の文書データと前記要約情報とのリンク結果に基づいて、検索された前記分析対象の文書データに対応する前記要約情報を検索し、

検索された前記分析対象の文書データと所属するカテゴリと検索された前記要約情報とを含む画面データを生成する。

According to a still another embodiment of the present invention, there is provided a system of document analysis comprising

- 20 a unit that 要約に含めるために抽出する要素として定義された要約要素を含む用語定義辞書データを参照し、分析対象の文書データに含まれている要約要素を抽出する、

a unit that 抽出された前記要約要素を所定の規則にしたがって組み合わせて前記分析対象の文書データの要約情報を生成する、 and

- 25 a unit that 前記分析対象の文書データと前記要約情報をリンクする。

According to a still further embodiment of the present invention, there is provided a system of document analysis comprising

a unit that 要約に含めるために抽出する要素として定義された要約要素を含む用語定義辞書データを参照し、分析対象の文書データに含まれている要約要素

を抽出する、

a unit that 抽出された前記要約要素を所定の規則にしたがって組み合わせて前記分析対象の文書データの要約情報を生成する、

a unit that 前記分析対象の文書データと前記要約情報をリンクする、 and

a unit that ユーザから前記要約情報の指定を受け付けた場合に、前記分析対象の文書データと前記要約情報とのリンク結果に基づいて、指定された前記要約情報に対応する前記分析対象の文書データを検索し、指定された前記要約情報と検索された前記分析対象の文書データとを含む画面データを生成する。

Additional objects and advantages of the invention will be set forth in the description which follows, and in part will be obvious from the description, or may be learned by practice of the invention. The objects and advantages of the invention may be realized and obtained by means of the instrumentalities and combinations particularly pointed out hereinbefore.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWING

The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate embodiments of the present invention and, together with the general description given above and the detailed description of the embodiments given below, serve to explain the principles of the present invention in which:

図 1 は、本発明の第 1 の実施の形態に係る文書分析システムの構成の一例を示すブロック図；

図 2 は、同実施の形態に係る文書分析システムにより生成される画面データの一例を示す図；

図 3 は、同実施の形態に係る文書分析システムの動作の一例を示すフローチャート；

図 4 は、要約抽出機能の抽出機能による要約要素の抽出結果の一例を示す図；

図 5 は、階層に基づいて表示条件を指定する状態の一例を示す図；

図 6 は、ユーザによって同階層の条件が指定される状態の一例を示す図；

図 7 は、同一階層の表示条件の指定を実現するための処理の一例を示すフローチャート；

図 8 は、過去の表示条件の指定と新規の表示条件の指定との組み合わせ方法の一例を示す図；

図 9 は、要約情報の指定により文書データの該当部分が強調表示された状態の一例を示す図；

図 10 は、文書分析プログラムにより実施されるサービスの提供形態の一例を示すブロック図。

DETAILED DESCRIPTION OF THE INVENTION

以下、図面を参照しながら本発明の実施の形態について説明する。なお、以下に示す各図において、同一の部分については同一の符号を付する。

(第 1 の実施の形態)

本実施の形態においては、報告の記載された文書データの分析作業を支援する文書分析システムについて説明する。

図 1 は、本実施の形態に係る文書分析支援システムの構成の一例を示すブロック図である。

文書分析システム 1 は、記録媒体 12 に記録されている文書分析プログラム 17 を読み出して実行する。

文書分析プログラム 17 は、システム 1 に読み込まれ、実行されると、取得機能 2、要約生成機能 3、操作受付機能 4、画面生成機能 5 を実現する。文書分析システム 1 は、データベース 6 に記録されている用語定義辞書 6a を参照する。

取得機能 2 は、分析対象の文書データを取得する。本実施の形態では、文書データは、メーカーの営業日報等の報告データであるとする。また、この文書データには、報告者名、報告日時、販売店名、日付などの文書データを分類するためのインデックス情報が含まれている。例えば、インデックス情報として、文書データの書誌事項を用いることができる。

要約に含めるために文書データから抽出する要素として定義された要約要素とその属性とは、互いに関連付けされた状態で用語定義辞書 6a に登録される。ユ

ーザは、要約要素として、例えば単語の一部、単語、句、節、表現など、抽出を望む内容を自由に定義できる。

例えば、用語定義辞書 6 a では、要約要素「スナック食品 A」に属性「自社商品」、要約要素「スナック食品 B」に属性「他社商品」が関連付けされているとする。また、要約要素「売れている」に属性「結果－優勢情報」、要約要素「販売不振です」に属性「結果－劣勢情報」が関連付けされているとする。また、要約要素「試食会」に属性「アクション」、要約要素「宣伝」に属性「アクション」が関連付けされているとする。

要約生成機能 3 は、抽出機能 7、分析機能 8、リンク機能 18 とを具備する。

抽出機能 7 は、取得機能 2 によって取得された文書データを受け付け、用語定義辞書 6 a を参照する。そして、抽出機能 7 は、用語定義辞書 6 a に登録されている要約要素と文書データとを比較し、用語定義辞書 6 a に登録されている要約要素と同じ表現が文章データ内に存在する場合に、その要約要素とその属性と文章データ中の位置情報を記録する。

分析機能 8 は、抽出機能 7 によって抽出された要約要素又は属性を所定の規則に基づいて組み合わせる要約情報を生成する。分析機能 8 には、例えば、抽出された要約要素を規則「商品－アクション」、規則「商品－結果」、規則「商品－アクション－結果」などにしたがって組み合わせる旨が設定されている。

分析機能 8 は、要約要素間を組み合わせてもよく、要約要素と属性を組み合わせてもよく、属性間を組み合わせてもよい。

抽出された要約情報又は属性の組み合わせ判定処理としては、例えば、アンド検索処理 8 a、文書区切り処理 8 b、係り受け分析処理 8 c、対応付け分析処理 8 d などがある。

操作受付機能 4 は、ユーザから判定処理の指定を受け付け、分析機能 8 に通知する。

アンド検索処理 8 a では、規則にしたがって抽出された要約要素又は属性の全ての組み合わせを生成する。

文書区切り処理 8 b では、所定の文書分割規則にしたがって文書データを区切り、この区切られた状態を使用して抽出機能 7 による抽出結果を組み合わせる。

例えば、文章データを「.」、「,」などで区切る。そして、区切られた範囲内で、抽出された要約要素又は属性を所定の規則にしたがって組み合わせる。

係り受け分析処理 8 c では、抽出された要約要素が比較の対象であるか否かを判定し、比較の対象であると判定された要約要素を組み合わせの候補から排除し、残りの要約要素を使用してアンド検索処理 8 a 又は文書区切り処理 8 b を実行する。例えば、比較の対象か否かは、「～er」「than」「far～than」「～と比べて」「対～比」などのような比較を意味する要素と、抽出された要約要素の位置とに基づいて判定される。

対応付け分析処理 8 d では、比較される関係にある要約要素を関連付けた対応テーブル 9 を参照する。また、対応付け分析処理 8 d では、抽出された要約要素に比較を意味する要素が付されておりこの要約要素の比較の対象となる要約要素が抽出されていない場合に、抽出された要約要素と比較される関係にある要約要素を対応テーブル 9 から求める。そして、対応付け分析処理 8 d では、抽出機能 7 によって抽出された要約要素と対応テーブル 9 から求められた要約要素とを組み合わせる。

例えば、対応テーブル 9 には互いに競合関係にある自社商品と他社商品とが対応付けされる。そして、「他社商品より売れている」という文書データについて分析を行ったとする。

この場合、比較を意味する「より」の付されている「他社商品」が抽出され、「他社商品」の比較対象がないため、対応テーブル 9 から「自社商品」が求められ、「自社商品」「売れている」という組み合わせ結果が得られる。

要約生成機能 3 は、例えば「…スナック食品 A はマーケットで 7 月に売れている…」という文書データについて、「スナック食品 A 売れている」などの要約情報を生成する。また、要約要素の属性から、この文書データに自社商品の優勢情報が含まれていることが把握される。

操作受付機能 4 は、要約生成機能 3 で使用する判定処理 8 a ～ 8 d の選定内容をユーザから受け付けると、その内容を要約生成機能 3 の分析機能 8 に通知する。

また、操作受付機能 4 は、画面表示に関するユーザの指定内容を受け付けると、その指定内容を画面生成機能 5 に通知する。

リンク機能18は、文書データと分析機能8によって生成された要約情報との間にリンクを設ける。リンク機能18は、同一の要約情報を持つ文書データを、この同一の要約情報を介してリンクで結合する。

画面生成機能5は、インデックス情報、要約生成機能3によって抽出された要約情報、日報本文である文書データを組み合わせた画面データを生成する。画面データは、ディスプレイ10に表示される。

図2は、文書分析支援システム1により生成される画面データの一例を示す図である。

画面11は、「期間」、「商品名称」、「業態」、「優勢情報か劣勢情報かの別」、「要約情報の内容」という順序の階層にしたがってユーザが表示条件を選択するための条件指定領域11a、11bを含む。なお、要約情報の内容を選択するための条件指定領域11bには、要約情報の内容毎に、その要約情報の抽出された文書データの件数が表示される。

表示条件は、インデックス情報及び要約情報とを階層的に組み合わせて指定される。

画面11は、表示条件の現在の指定状態を表示する領域11cを含む。

画面11は、指定された表示条件を満たす文書データ、その文書データから生成された全ての要約情報、その文書データの含むインデックス情報とを組み合わせて一覧表示する一覧領域11dを含む。

画面11を参照したユーザが一覧領域11dに掲載されているあるインデックス情報を操作受付機能4により指定した場合、画面生成機能5は、指定されたインデックス情報を含む文書データを検索する。

そして、画面生成機能5は、検索された文書データ、検索された文書データに含まれているインデックス情報、検索された文書データから生成された要約情報とを組み合わせ、一覧表示する画面データを生成する。

一方、画面11を参照したユーザが一覧領域11dに掲載されているある要約情報を操作受付機能4により指定した場合、画面生成機能5は、指定された要約情報とリンクされている文書データを検索する。そして、検索された文書データ、検索された文書データに含まれているインデックス情報、検索された文書データ

から生成された要約情報とを組み合わせ、一覧表示する画面データを生成する。

5 このように、画面生成機能 5 は、ユーザに指定された要約情報又はインデックス情報にしたがって文書データを検索する情報検索処理 5 a、ユーザが階層的に指定した表示条件（検索キー）にしたがって文書データを検索する階層検索処理 5 b を具備する。

画面生成機能 5 は、文書データのうち要約情報に対応する部分の表示性質を変更する表示性質変更処理 5 c、検索された文書データを XML（Extensible Markup Language）で記述する構造化処理 5 d を具備する。

10 図 3 は、上記のような構成を持つ文書分析システム 1 の動作の一例を示すフローチャートである。

ステップ S 1 において、文書分析システム 1 の取得機能 2 は、分析対象の文書データを読み込む。

ステップ S 2 において、文書分析システム 1 の抽出機能 7 は、読み込んだ各文書データから所定の要約要素を抽出する。

15 ステップ S 3 において、文書分析システム 1 の分析機能 8 は、抽出された要約要素に基づいて要約情報を生成する。

ステップ S 4 において、文書分析システム 1 のリンク機能 1 8 は、文書データと要約情報とをリンクする。

20 ステップ S 5 において、文書分析システム 1 の画面生成機能 5 は、ユーザが表示条件を指定するための条件指定領域 1 1 a、1 1 b を含む画面 1 1 を表示する。

ユーザは、条件指定領域 1 1 a におけるプルダウンメニュー又は条件指定領域 1 1 b のリストを使用して表示する文書データを指定する。

25 例えば、ユーザは、インデックス情報の日付が「2001 年 3 月 1 日～2001 年 3 月 31 日」内であり、インデックス情報の商品が「スナック食品 A」「スナック食品 B」であり、要約情報が属性「優勢情報」である旨を示し、要約情報「景品を付けるので売行きがよい」にリンクされていることを表示条件として指定する。

ステップ S 6 において、文書分析システム 1 の操作受付機能 4 は、ユーザの指定した表示条件を受け付ける。

ステップS7において、画面生成機能5は、この表示条件を満たす文書データとその要約情報とそのインデックス情報とを組み合わせた一覧を表示する。

ステップS8において、文書分析システム1は、ユーザによる分析作業が継続する限り、表示条件の指定の受け付けとこの表示条件を満たす内容の表示を繰り返す。ユーザは、一覧に表示されているインデックス情報と要約情報とを参照し、さらに分析を継続したい場合に、インデックス情報又は要約情報の表示をマウスで指定（クリック）し、新たな表示条件を指定する。インデックス情報と要約情報とは、自由に組み合わせて表示条件として指定できる。

上述したように、文書分析支援プログラム1は、このユーザの指定した表示条件を受け付け、この表示条件を満たす文書データとその要約情報とそのインデックス情報とを組み合わせた新規の一覧を表示する。

以下に、文書分析システム1を使用することにより得られる効果について説明する。

例えば、企業で使われる日報データ、月報データ、営業報告データ、店頭管理日報データ等の文書データの数は膨大になる。

ユーザは、文書分析システム1を起動し、収集された文書データをこの文書分析システム1に読み込ませる。すると、文書データに基づいて要約情報が生成される。

ユーザは、文書分析システム1を使用し、生成された要約情報の内容にしたがって、文書データを分類し、集計すると、「ある商品に関する情報が多い」、「ある販売促進活動をしたところ売れている」という情報が多い、「ある競合他社商品に関する情報が多い」などのような定量的な情報を容易に求めることができる。

そして、ユーザは、商品別、メーカー別、営業部署別などの観点で文書データを自動的に分類し、分析に利用できる。

要約情報別の件数を表示することで、ユーザは、検索などを実行しなくても市場状況を把握できる。

ユーザは、表示された要約情報を読むことで大量の文書データ全てを読まなくても内容を把握することができる。

文書分析システム 1 では、表示条件がユーザに指定された場合、上記図 2 の画面 1 1 に示すように、ユーザに指定された表示条件と異なる意味の表示条件が検索結果とともに表示される。

具体的には、要約情報「景品を付けるので売行きがよい」という表示条件が指定された場合、指定された表示条件に基づいて検索された文書データのみではなく、例えば「ラッピングをするが売行きが悪い」などのような検索された文書データにリンクされており指定された要約情報と全く異なる他の要約情報までも表示される。インデックス情報についても同様である。

単に、ユーザに表示条件を階層的に指定させるとすると、「劣勢情報」の中から「ラッピングをするが売行きが悪い」という表示条件を指定するためには、ユーザは、まず「劣勢情報」を指定し、次に「ラッピングをするが売行きが悪い」を指定する必要がある。しかしながら、文書分析システム 1 では、表示条件を階層的に指定するのみではなく、ある表示条件に基づいて表示された画面から他の表示条件に基づいて表示される画面に直接移行する機能が具備されており、ユーザの操作性が向上する。

すなわち、文書分析システム 1 を利用すれば、ある条件を満たす一覧から他の条件を満たす一覧に柔軟に移行することができる。また、文書分析システム 1 を利用すれば、ユーザは表示条件を階層に関係なく適宜指定することができるため、ユーザの操作性を向上させることができる。

(第 2 の実施の形態)

本実施の形態においては、上記第 1 の実施の形態における要約生成機能 3 について詳細に説明する。

用語定義辞書 6 a には、「スナック食品 A」「スナック食品 B」「スナック食品 C」のような商品名の要約要素と、「試食会」「完売」「売れている」のようなアクション又は結果に関する要約要素とが登録されているとする。

また、要約生成機能 3 の抽出機能 7 には、「スナック食品 B は試食会で完売。スナック食品 A の情報。スナック食品 C に対して 1 2 0 % で売れている。」という文章データが受け付けられたとする。

この場合、抽出機能 7 は、この文書データと用語定義辞書 6 との双方に含まれ

ている商品名の要約要素「スナック食品A」「スナック食品B」「スナック食品C」と、アクション又は結果に関する要約要素「試食会」「完売」「売れている」を抽出する。

図4は、要約生成機能3の抽出機能7による要約要素の抽出結果の一例を示す図である。要約要素とその位置とその要素IDとが抽出されている。

要約生成機能3の分析機能8は、抽出された要約要素を所定の規則に基づいて組み合わせて要約情報を生成する。

対応テーブル9は、対応付け分析処理8dに参照されるテーブルである。この対応テーブル9には、競合関係にある商品名「スナック食品A」「スナック食品B」「スナック食品C」が関連付けられて登録されている。

上記の文書データ「スナック食品Bは試食会で完売。スナック食品Aの情報。スナック食品Cに対して120%で売れている。」について「商品」と「アクション又は結果」の正しい組み合わせは、「スナック食品B－試食会」「スナック食品B－完売」「スナック食品A－売れている」の3つである。

上記各種判定処理8a～8dの分析精度を適合率（生成された要約の内容が正しい割合）及び再現率（要約に含まれていなければならない内容が実際に含まれている割合）で評価すると以下になる。なお、組み合わせ規則は「商品－アクション」、「商品－結果」とする。

アンド検索処理8aでは、抽出された要約要素の全ての組み合わせを規則にしたがって生成する。したがって、このアンド検索処理8aで生成される要約情報は「スナック食品B－試食会」「スナック食品B－完売」「スナック食品B－売れている」「スナック食品A－試食会」「スナック食品A－完売」「スナック食品A－売れている」「スナック食品C－試食会」「スナック食品C－完売」「スナック食品C－売れている」の9つである。この結果の適合率は約33%、再現率は100%である。したがって、ユーザは、文書データから再現率を優先して要約情報を生成する場合、操作受付機能4でアンド検索処理8aを選択する。

文書区切り処理8bでは、「。」で文書データを区切り、その範囲内でアンド検索を行う。したがって、この文書区切り処理8bで生成される要約情報は「スナック食品B－試食会」「スナック食品B－完売」「スナック食品C－売れてい

る」の3つである。この結果の適合率は約66%、再現率は約66%となる。したがって、ユーザは、文書データから適合率と再現率とを同程度に優先して要約情報を生成する場合、操作受付機能4で文書区切り処理8bを選択する。

係り受け分析処理8cは、「。」で区切られた範囲内又はその範囲より前に位置し、該当要素にもっとも近い位置にあり、組み合わせから除外する判断基準として定義されている所定の除外用語に関係しない商品を探し、組み合わせを行う。したがって、この係り受け分析処理8cで抽出される要約情報は「スナック食品B-試食会」「スナック食品B-完売」「スナック食品A-売れている」の3つである。この結果の適合率は適合率100%、再現率は100%となる。

対応付け分析処理8dは、係り受け分析処理で商品が抽出されなかった場合に、除外用語に関係する他社商品に対応する自社商品に対応テーブル9から求め、求めた自社商品を使用して組み合わせを行う。したがって、この対応付け分析処理8dで生成される要約情報は「スナック食品B-試食会」「スナック食品B-完売」「スナック食品A-売れている」の3つである。この結果の適合率は100%、再現率は100%となる。

したがって、ユーザは、文書データから適合率と再現率の双方を優先して要約情報を生成する場合、操作受付機能4で係り受け分析処理8c又は対応付け分析処理8dを選択する。

そして、要約生成機能3は、「自社商品」に対して優勢情報の結果又は優勢情報のアクションが組み合わせられている場合には、この要約情報を優勢情報であると判定する。

一方、要約生成機能3は、「自社商品」に対して劣勢情報の結果又は劣勢情報のアクションが組み合わせられている場合、及び「他社商品」に対して優勢情報の結果又は優勢情報のアクションが組み合わせられている場合、この要約情報を劣勢情報であると判定する。

以上説明したように、文書分析システム1では、要約情報を生成する分析機能8において複数の判定処理8a~8dを実行可能であり、ユーザはこの判定処理8a~8dの自由に選択できる。したがって、分析対象の文書データの質やユーザのニーズに合わせて柔軟に表示を変化させることができる。

(第3の実施の形態)

本実施の形態においては、上記第1の実施の形態に係る文書分析システム1の変形例について説明する。

図5は、階層に基づいて表示条件を指定する状態の一例を示す図である。この図5では、まず第1階層でメーカーについての表示条件が指定され、次に第2階層でそのメーカーの商品についての表示条件が指定される。

このように、ユーザによって指定された表示条件よりも下位の表示条件が指定される方式では、同じ階層の表示条件が複数指定されることはない。例えば、メーカーM1とメーカーM2の双方を指定することはできない。

したがって、「メーカーM2のスナック食品BとメーカーM3のスナック食品Cの情報とを含む文書データを表示する」というニーズがあっても、M2社のスナック食品Bに関する文書データの中からユーザ自身がM3社のスナック食品Cに関する文書データを抽出するか、またはM3社のスナック食品Cに関する文書データの中からユーザ自身がM2社のスナック食品Bに関する文書データを抽出するしか手段はない。

そこで、本実施の形態における画面生成機能5では、図6に示すように、指定された表示条件と同階層の表示条件をユーザが指定可能とするために、例えばメーカーM1、M2などのように、同じ階層レベルにある表示条件を上下の階層として指定可能とする。

図6は、ユーザによって同階層の条件が指定される状態の一例を示す図である。本実施の形態における画面生成機能5は、ユーザがある表示条件を指定すると、その指定された表示条件と上下関係にある一つ下の階層における全ての表示条件に加えて、指定された表示条件と同じ階層に属し未指定の表示条件を含むリスト、及び「文書表示」を表示する。

そして、画面生成機能5は、ユーザによって「文書表示」が指定された段階で指定された表示条件を満たす文書データとその要約情報とそのインデックス情報とを検索し、組み合わせて画面データを生成する。

図6では、まず、全てのメーカーM1～Mmの名称が表示条件のリストとして表示されている。ユーザがこのリストから「メーカーM2」を指定すると、メー

カーM2の商品「商品P1」～「Pp」及びメーカーM2を除く「メーカーM1」、「メーカーM3」～「メーカーMm」の記載されたリストが表示される。

図7は、同一階層の表示条件の指定を実現するための処理の一例を示すフローチャートである。

5 ステップT1において、画面生成機能5は、ある階層の表示条件と「文書表示」の記載されたリストを表示する。

ステップT2において、画面生成機能5は、リストに対する指示を受け付ける。

ステップT3において、画面生成機能5は、「文書表示」が指定されるか否かを判断する。

10 「文書表示」が指定されなかった場合、ステップT4において、画面生成機能5は、「最新指定」フラグの付された表示条件のフラグを「指定」フラグに変更する。

ステップT5において、画面生成機能5は、新たに指定された表示条件に「最新指定」フラグを付す。

15 ステップT6において、画面生成機能5は、「最新指定」フラグの付された表示条件の下位の表示条件と、「最新指定」フラグの付された表示条件と同一階層でありフラグの付されていない表示条件と、「文書表示」とを記載したリストを表示する。

「文書表示」が指定されるまでステップT2以下の処理が繰り返される。「文書表示」が指定されると、ステップT7において、画面生成機能5は、「指定」フラグの付された全ての表示条件を検索キーとして文書データを検索し、画面データを生成する。

20 本実施の形態においては、ユーザが同じ階層の複数の表示条件を指定可能とした。これにより、例えば「メーカー名」「要約情報」「文書データ」というトップダウンの表示条件の指定のみではなく、同じ階層の表示条件を柔軟に指定できるため、ユーザの操作性を向上させることができる。したがって、例えば「メーカー」「要約情報」「文書データ」というように表示条件の階層と階層の数とが固定的に定められている場合よりも、ユーザのニーズに応じた検索が可能となる。

25 なお、本実施の形態においては、「メーカー」について同一階層の指定を可能

とする場合について説明しているが、他の階層においても同一階層で複数の表示条件を指定可能にできる。また、複数の階層において同一階層の表示条件を指定可能としてもよい。

（第4の実施の形態）

5 本実施の形態においては、上記第3の実施の形態に係る文書分析システム1の変形例について説明する。

本実施の形態においては、上記各実施の形態と同様に、表示されている文書データと要約情報との間でリンクを設けている。そして、例えば「ラッピングをするが発行きが悪い」という要約情報をクリックすると、この要約情報とリンクしている文書データを画面に表示する。本実施の形態における画面移行には、上記
10 第3の実施の形態で述べた表示条件の指定方法を利用する。

図8は、過去の表示条件の指定と新規の表示条件の指定との組み合わせ方法の一例を示す図である。

ユーザは、「メーカーM2」「メーカーM1」「文書表示」と表示条件を絞り
15 込んだとする。この場合、この表示条件を満たす文書データが検索され、画面19が表示される。

表示された画面19上で新規に「メーカーM1」「商品P2」が表示条件として指定されたとする。この場合、画面生成機能5は、図8の実線の矢印で示したように、過去に指定されたユーザの絞り込み順序を逆にたどり、「メーカーM
20 1」が指定された状態に戻り、この「メーカーM1」の下位の表示条件として「商品P2」が指定されたとする。

本実施の形態においては、新規な表示条件と同じ指定をユーザが過去に行っていた場合又は新規な表示条件と同じ階層に属する表示条件が過去にユーザに指定されていた場合には、その過去に指定された表示条件に至るまでの表示条件と新規な表示条件とを含めて表示条件が構成され、文書データが検索される。
25

一方、表示された画面上で未だユーザによって指定されていない表示条件が指定された場合には、階層のトップに戻り、指定された表示条件のみに基づいて文書データが検索される。

したがって、ユーザは、過去の絞り込み操作を活かしたまま表示条件を指定し、

文書データを表示できる。これにより、ユーザは、容易に限定した表示内容を得ることができる。

(第5の実施の形態)

本実施の形態においては、上記第1から第4までの実施の形態に係る文書分析システム1の変形例について説明する。

画面生成機能5は、要約情報がクリックされると、文書データのうちこの要約情報に対応する部分を強調表示する。

図9は、要約情報の表示欄で要約情報「景品をつけた」がクリックされ、文書データの該当する部分「景品をつけて」が強調表示されている。

このような機能は、要約生成機能3による要約情報生成時に、文書データ中に要約情報の生成結果をタグとして挿入しておき、要約情報欄の記載と対応付けることで実現可能である。

例えば、HTMLファイルの場合、要約情報とそれに対応する文書データの記載との間をリンクで結合し、クリックされた場合に該当部分が強調表示されたHTMLファイルを表示する。

なお、例えば、予め要約情報をその要約情報の種別に応じた色で表示し、その要約情報に対応する文書データの記載を、その要約情報の種別に応じた色で表示してもよい。

このように文書データから生成された要約情報が文書データ中のどこに対応するかをユーザに明確に示すことで、文書データの記載量が多くてもユーザは要約情報の具体的な記載内容を素早く知ることができる。

また、ユーザは要約情報を持つ文書データ全文を読まなくても要約情報に該当する記載の前後を読んで内容を把握できるため、情報集約度を高くすることができる。

(第6の実施の形態)

本実施の形態においては、上記第1から第5までの実施の形態に係る文書分析システム1の変形例について説明する。

画面生成機能5は、画面の文書データの表示部分をXMLにて記述する。その結果、上記各実施の形態のように複数の文書データを容易に組み合わせることが

可能である。

画面の文書データの表示部分をXMLで記述することで、複数の文書データを含む電子ファイルから各文書データを任意に選択し、組み合わせることができる。

ユーザは、検索された文書データをさらに編集可能であり、ユーザがさらに情報を集約して関係者に報告することが可能になり、ナレッジマネジメントシステムとしての利便性が向上する。

なお、上記各実施の形態に係る文書分析システム1によって実現される機能は、同様の作用・機能を実現可能であれば配置を変更させてもよく、また各機能を自由に組み合わせてもよい。

また、上記各実施の形態における文書分析プログラム17によって実現される機能2～5は、複数のコンピュータ上に分散され、互いに連携しつつ動作してもよい。

上記各実施の形態で説明した文書分析プログラム17は、例えば磁気ディスク（フレキシブルディスク、ハードディスク等）、光ディスク（CD-ROM、DVD等）、半導体メモリなどの記録媒体12に書き込んでコンピュータに適用可能である。また、このプログラムは通信媒体により伝送して、計算機、計算機システムに適用可能である。

コンピュータは、記録媒体12に記録された文書分析プログラム17を記録媒体12から読み込み、プログラムによって動作が制御されることにより、上述した機能を実現する。

（第7の実施の形態）

本実施の形態においては、上記各実施の形態で説明した文書分析プログラム17の利用態様について説明する。

図10は、上記各実施の形態で説明した文書分析プログラム17により実施されるサービスをASP（アプリケーション・サービス・プロバイダ）が提供する形態の一例を示すブロック図である。

ユーザ13は、自己の端末14から例えばインターネットなどのようなネットワーク15を経由してASP16の管理する文書分析プログラム17を利用する。これにより、文書データの分析作業を効率的にかつ容易に実施できる。

また、ASP16のサービスの提供を受けることで、ユーザ13は、自己で文書分析プログラム17を管理する場合よりも保守、運用の面で効率的に分析サービスを利用できる。

ASP16は、ユーザに分析支援サービスを提供し、ユーザから対価を得ることが
5 できる。

While the description above refers to particular embodiments of the present invention, it will be understood that many modifications may be made without departing from the spirit thereof. The accompanying
10 claims are intended to cover such modifications as would fall within the true scope and spirit of the present invention. The presently disclosed embodiments are therefore to be considered in all respects as illustrative and not restrictive, the scope of the invention being indicated by the appended claims, rather than the foregoing description, and all changes that come within the meaning and range of equivalency of
15 the claims are therefore intended to be embraced therein.